

DESCRIPTION

TITLE

“Organiser for complex categorisations”

5

TECHNICAL AREA

The present invention falls within the area of computerized tools for facilitating the classification of information.

10 PRIOR ART

In the present document the following references are cited:

[1] Amazon. Book browser. www.amazon.com.

[2] Barnes and Noble. Book browser. www.bn.com.

[3] Benson, J.D., Cummings, M., Greaves, W.S. (eds) (1988) “Linguistics in a Systemic
15 Perspective”, Amsterdam: John Benjamins Publishing Company

[4] IBM (2000). US Patent. N. 6.055.515

[5] Microsoft. MSDN Library Visual Studio 6.0

[6] Royal Academy of Spanish Language. Dictionary of Spanish Language. Espasa.

20 The appearance of informatics in general and of Internet in particular have caused that there nowadays exists a growing amount of information that is available for a great amount of persons, many of whom are not expert user of informatics. For example, there nowadays exists a great variety of databases that are accessible in CD-ROM, DVD or in Internet servers. Some examples of these databases are the following ones:

- 25 1. The electronic dictionary of the Royal Academy of Spanish Language.
2. The encyclopedia Microsoft® Encarta®.
3. The help tool in Microsoft® Visual Studio®.
4. The topic hierarchy in Yahoo®.
5. The book catalogs in Amazon®, Barnes and Noble® and other online bookstores.

30 In general, these databases are organized in such a way that they contain a mixture of many different types of concepts, and searches are difficult to perform. This creates a need for new and better mechanisms for organizing the information and facilitating the execution of searches.

35

EXPLANATION OF THE INVENTION

ESENCE OF THE INVENTION

This invention presents an approach for organizing sets of entities, wherein said entities
 5 might be for example objects, concepts, ideas, terms or others, which facilitates the conceptualization of classifications and the execution of searches. In particular, the invention facilitates the creation of systematic categorizations in which there exist different criteria for organizing the information, and aids the user in the inspection and utilization of such classifications.

10 The invention unites in the same tree the categories that are used for classifying the instances that are being classified and the different criteria that define the different hierarchies of categories. That is to say, it creates a multicriteria classification in the same tree, in which coexist different hierarchies that belong to different criteria.

This tree can be graphically shown in an arboreal structure. Exhibit 1 shows a simple
 15 example of a possible arboreal structure for a multicriteria classification of words. In order to facilitate the exposition, in this document the term ‘tree’ will be used for the logical organization of entities by means of parent-child relationships, and the term ‘arboreal structure’ will be used for the representation of said tree in a graphical interface.

Also with the purpose of facilitating the exposition, the following guidelines will be
 20 followed in the different arboreal structures that will be shown in the document:

- the different instances will be written between dots, such as for example “.hammer.” —in the example of Exhibit 1, the instances would be the example words which are shown,
- the different categories that will be used will be shown with a normal font characteristics, such as for example “Noun” in the example of Exhibit 1, and
- 25 – the different criteria that will be used will be shown with underlined font, such as for example “According to nature”, in the example of Exhibit 1.

It must be taken into account that the purpose of the categorization shown in Exhibit 1, and of other categorizations that will be shown below, is only to facilitate the explanation of the
 30 invention, and that the particular decisions that will be taken for criteria, categories or instances are only intended as examples, and not to limit in any way the scope of the invention.

Exhibit 1

Words

35 Noun

According to nature

Entity

.hammer.

.brother.

5 .writer.

.cherry.

Attribute

.height.

.honesty.

10 Event

According to duration

Punctual

.arrival.

Durative

15 .concert.

.storm.

According to action

Action

.concert.

20 .arrival.

No action

.storm.

Other

.meter.

25 .field.

According to meaning

Has utilization

.hammer.

Has function

30 .writer.

Has relationship

.brother.

Other

.height.

35 Verb

Adjective

Adverb

Closed Class

5 As can be seen, this approach allows to mix different categories and different criteria in an unlimited way. That is to say, both categories and criteria can be descendent nodes of a criterion. And both categories and criteria can be descendent of a category. In the example shown in Exhibit 1, there is no case in which a criterion is a child of another criterion, but this might happen as well.

10 As can be seen, the categorization shown in Exhibit 1 can be built in a simple way in a single control of the types used for showing trees; it could be shown, for example, in a tree control of those which are customarily used, such as the Microsoft Treeview® control, which is used in the directory structure of the operating system Microsoft Windows®. In this case, both criteria and categories could be implemented as nodes of the arboreal structure that the control
15 represents.

 As can also be seen, in order to facilitate the utilization of the invention, in general, there might optionally exist graphical means that would allow to distinguish criterion-nodes from category-nodes, such as has been done in Exhibit 1, for example, using underlined font for criterion-nodes, even though it would be possible to use other type of means.

20 As can also be seen, instances can belong to different categories; in particular, they will normally belong to different categories that are descendents of sister criterion-nodes. For example, in Exhibit 1, “.arrival.” belongs to “According to duration > Punctual” and it also belongs to “According to Action > Action”. Depending on the criterion that is taken into account, the word will belong to different categories. As is customarily done in the prior art, the
25 arboreal structure can show the different instances in a duplicated or repeated fashion in different positions that correspond to different categories. That is to say, “.arrival.” is located both in “According to duration > Punctual” as in “According to Action > Action”.

 Finally, the invention applies both to a case in which no instances exist yet, and therefore only criteria and categories appear, and to the case in which there exist instances that
30 are shown.

 Moreover, the invention can be used in a case in which only criteria and categories are shown and instances are not shown. In this case, categories and criteria could be used to execute searches against a database in which the instances are stored.

OPTIONAL FEATURES

The invention allows to implement different embodiments that have different optional features. In order to facilitate the explanation of the advantages of the invention, and without any limiting intention, some of these optional features are explained below:

5

1. A relational database is used which contains two tables. One table is used for storing the instances and the other table is used for storing categories and criteria.
2. A different code is assigned to each record of the category-criterion table (i.e., each category and each criterion have a different code). For example, a numeric code is assigned where
10 each code is an integer number.
3. A special field is created in the instance table, called for example “Classification”,
4. For each instance, the categories to which the instance belongs are identified and the codes of those categories are concatenated, creating a string, and some delimiting character is used in both sides each of the codes. In a hypothetical case, the linked string could be a string such as
15 this one “-1-23-42-100-230-”.
5. The linked string for the codes of each instance is stored in the field “Classification” of each instance.

With the previous method, it would be possible to undertake very powerful searches on
20 the instances. It could be possible for example to execute queries on the instances that have certain categories by using the SQL command “Like” or another similar command in other database language. For example, if the user wants to find instances that are assigned to the category whose code is “1”, the query condition could be “Classification LIKE ‘-1-’”. This condition would retrieve all the instances that would have the term “1” in its classification. As
25 can be seen, the utilization of delimiting characters prevents wrong results from being retrieved, such as for example would happen with the code “100”; without the delimiting characters, it might happen that this wrong code would be retrieved, because the command “LIKE” would consider that “100” contains a character “1”.

There is another optional feature that is helpful for executing queries, and which is
30 characterized by the fact that the user can select a set of categories, which can belong or not to different classification criteria, and the system would search for those instances that have certain relationship with those categories. For example, for the data in Exhibit 1, the user could select only the category “Entity” (which is hanging from Noun > According to Nature) and the system would return the instances: “.hammer.”, “.brother.”, “.writer.”, and “.cherry.”, i.e. all the
35 instances that belong to the category “Entity”. Alternatively, if the user selects simultaneously

“Entity” and “Has utilization” (where “Has utilization” is hanging from Noun > According to meaning), the system would only return “.hammer.”, because it is the only instance that belongs to both categories. The user can perform queries as complex as he/she wishes by using boolean expressions in order to refine the conditions that must be imposed over the categories that are
 5 selected. For example, if the user selects “Entity” and NO “Has utilization” (where “NO” is the boolean function ‘negation’) the system would return only “.brother.”, “.writer.” and “.cherry.”.

In order to even more ease query formulation, a useful optional feature is the SUMMARY ARBOREAL STRUCTURE. A summary arboreal structure is an arboreal structure
 10 that only contains the nodes that are selected in the main arboreal structure at a given moment. For example, in Exhibit 1 it is possible to select certain nodes, such as for example the ones that are shown in bold font in Exhibit 2. A possible summary arboreal structure for this structure would be the structure shown in Exhibit 3.

15 **Exhibit 2**

Words

Noun

According to nature

Entity

20 .hammer.
 .brother.
 .writer.
 .cherry.

Attribute

25 .height.
 .honesty.

Event

According to duration

Punctual

30 .arrival.

Durative

 .concert.
 .storm.

According to action

35 **Action**

	.concert.
	.storm.
	No action
	.storm.
5	Other
	.meter.
	.field.
	<u>According to meaning</u>
	Has utilization
10	.hammer.
	Has function
	.writer.
	Has relationship
	.brother.
15	Other
	.height.
	Verb
	Adjective
	Adverb
20	Closed Class

Exhibit 3

	Words
	Noun
25	<u>According to nature</u>
	Entity
	Event
	<u>According to action</u>
	Action
30	<u>According to meaning</u>
	Has relationship

In order to further ease the management of the selected nodes, the information of the summary arboreal structure could also be shown as appears in Exhibit 4.

Exhibit 4

Words > Noun > According to nature > **Entity**

Words > Noun > According to nature > Event > According to action > **Action**

Words > Noun > According to meaning > **Has relationship**

5

It is also possible to add a nickname to the selected nodes, in order to more easily use them in queries, as shown in Exhibit 5.

Exhibit 5

Node	NICK
Words > Noun > <u>According to nature</u> > Entity	Entity1
Words > Noun > <u>According to nature</u> > Event > <u>According to action</u> > Action	Action1
Words > Noun > <u>According to meaning</u> > Has relationship	Relation1

10

One more optional feature that can be implemented is related with query generation after a selection of instances. When selecting an instance, it is possible to automatically select the categories to which that instance belongs, and the user then can take those categories as the starting point to generate queries.

15

ADVANTAGES OF THE INVENTION

1. It allows to easily merge categorizations that are based on different criteria, so that the user can easily comprehend the effects of the multi-categorization.
2. It allows to easily perform sophisticated queries, because the user only has to select in the same control all the categories that he/she wishes, and then only has to combine them in order to create the query.
3. It allows to create simple user interfaces, because it allows to create multicriteria classifications with a single tree control.
4. It allows to flexibly create databases. If the user wants to change anything in the nodes of the tree, he/she only needs to add more records to the data base (in order to create additional nodes) and modify the field "Classification" for the instances, so it is not necessary to modify the structure of the database.
5. It allows to flexibly create user interfaces. If the user wants to change anything, such as for example adding or removing any criterion, it is not necessary to modify the programs that manage the arboreal structure nor modify the ones that manage the user interface, because the only thing that must be done is adding more nodes to the tree.

25
30

6. If facilitates the application of data mining systems, because different categories can be evaluated in an independent fashion. For the example of Exhibit 1, it is possible to analyze the word “.writer.” from the point of view of “Has Function” and from the point of view “Entity”.

5 The queries that are based on commands such as “LIKE” are relatively slow, which is a disadvantage. However, when the database is being created, which is when flexibility is more needed, there are normally few records, and therefore the effects of this disadvantage are smaller.

 If the user wants to increase the speed of the queries, he/she can modify the structure of the database, adding new fields for “Classification”, so that the different categories can be spread
10 over different fields, which would speed up the execution of searches. In these circumstances, it is possible to decide that certain fields will host the categories that might experience the least variation, and create one or more fields to host those linked codes that belong to the categories that can vary the most.

15

COMPARISON WITH OTHER PROPOSALS THAT EXIST IN THE PRIOR ART.

 As far as has been known, there do not exist proposals like this invention, even though some proposals share some features. The most similar proposals are the following ones:

20 Proposals in systemic linguistics. In this school of linguistic research many linguistic taxonomies are normally performed. A sample of papers on this area can be found in [Benson et al 1988]. In the tradition of systemic linguistics, linguistic entities are categorized by using diagrams that have some characteristics similar to those of the diagram shown in Figure 1, which has been taken from [Benson et al 1988, p.326]. In these diagrams it is possible to see that some parts
25 correspond to categories, and some other parts are similar to what in this invention is called criteria.

 However, despite the fact that this type of diagrams have been used since a long time ago (at least since 1988, which is the date of the reference), and despite the fact that there exist several computer tools to manage this type of diagrams, as far as has been known there are no
30 proposals that are similar to the one in this invention. The diagrams that are used in this tradition have the same format as a two dimensional picture, as shown in Figure 1, which is much more difficult to use than a tree control as the Microsoft TreeView control. For example, these diagrams do not have the selection possibilities that tree controls have, such as expanding and collapsing nodes. Furthermore, the diagrams expand from left to right and top down, which make
35 it difficult to manage the user interface. Moreover, it is necessary to create special purpose

computer programs in order to manage these diagrams. Additionally, there does not exist a clear distinction between criteria and categories. It has not been possible to find any proposal where all the diagram is integrated into a control in such a way that can be easily used.

In contrast, the invention described in this document can be easily implemented in a standard tree control, such as the Microsoft Treeview® control, or linking text strings in HTML language.

Classifications where diverse aspects appear in a mixed fashion. In this proposals, the classifications have some nodes that represent categories, other ones that might resemble criteria but which actually are not criteria, and other ones that represent additional aspects. Most of the proposals that have been found correspond to classifications where different aspects are mixed. In these proposals there is no method for classification and search that facilitates the user the utilization of the categorization. Nodes of different types appear in a mixed fashion, which creates confusion to the user. A selection of some proposals of this type is the following one: [Royal Academy of Spanish Language], [IBM 2000], [Microsoft]. [Amazon], [Barnes and Noble].

For example, in [Royal Academy of Spanish Language], it is possible to see a classification such as the one shown in Exhibit 6 (In the Exhibit, criteria and categories have been translated into English, but instances remain in Spanish). In this proposal, it is possible to select any of the existing categories in order to explore the instances that depend on that category. In the proposal, there do not really exist nodes that correspond to what in the present invention is called “criteria”. Even though some nodes might look like classification criteria, actually they are categories. For example, the adjective “.alto.” (“tall” in English) does not appear under the category-node “gender -> masculine”, but it is reserved for those adjectives that are only masculine, and they do not appear in the category “adjective”. The adjective “.altísimo.” (“very tall” in English) only appears in “levels->superlative”, and the adjective “.encinta.” (“pregnant”) only appears in “gender->feminine”. However, the word “.tanto.” (similar to “both”) appears both in “adjective” and in “uses as adjective”.

Exhibit 6

adjectives

adjective

.alto.

.tanto.

uses as adjective

.tanto.

	gender
	masculine
	.alto.
	feminine
5	.encinta.
	invariable
	levels
	comparative
	superlative
10	.altísimo.
	types
	anaphoric
	descriptive
	demonstrative
15	epithet
	gentilic
	indefinite
	possessive
	Latin adjectives
20	adjective locutions

In [Microsoft] it is possible to see a classification as the one shown in Exhibit 7 (many nodes have been omitted in order to facilitate the exposition). In this case the tree contains a great variety of topics, which are organized in the same way as the manner in which they could have appeared in a book that was structured in chapters and epigraphs, and it happens that some instances appear in several nodes, such as for example the control “CheckBox Control”. However, this is not a classification like the one proposed in this invention because, among other things, it does not have the criterion-node concept.

30 **Exhibit 7**

MSDN Library Visual Studio 6.0

 Welcome to the MSDN Library

 Visual Studio Documentation

 Visual Basic Documentation

35 Using Visual Basic

Reference

Language Reference

Objects

.CheckBox Control.

5

Properties

Controls Reference

Intrinsic Controls

.CheckBox Control.

10

In [IBM 2000] it is possible to see a classification like the one shown in Exhibit 8, which in [IBM 2000] is used as an example to propose an invention related to browsers for classifications.

Exhibit 8

15 Application

Accountancy

.ABC-123.

.XYZ-890.

.Programming.

20

.Typing.

Catalog

.Desktop Publishing.

Spreadsheet

.ABC-123.

25

.XYZ-890.

.Word processing.

Manufacturer

Company A

.ABC-123.

30

Company B

.XYZ-890.

Given that this classification is only a limited example, it is difficult to know exactly the intention of the authors of this patent. However, the most appropriate interpretation is that this

proposal is again a classification that mixes heterogeneous entities, such as in [Microsoft]. The reasons that explain this are the following ones:

1. The intention of the author is only to show a classification that contains multiple paths to an instance, independently of whether those multiple paths exist because there exist several criteria. An alternative classification in this respect might be the one shown in Exhibit 9, in which the product .XXX. might appear in two different nodes, and which however has no relationship with the current invention. The following sentences extracted from the patent show that that was the intention of the authors: “enter items that can be subcategories or products of several different categories”, and “a user should be able to navigate to a pair of sunglasses by following a path through many categories, such as beach wear, or sportswear of eye care”, wherein the situation that is described by this last sentence is represented in Exhibit 10.
2. In [IBM 2000] the discussion always mentions categories and products, and no distinction is made between category types, which shows clearly that the criterion concept does not appear in the text.
3. The classification that is shown only mixes different concepts, as is proven by the fact that there exist three root nodes, which do not depend on a single category node, in a similar way as how [Microsoft] links different concepts. If this classification had been taken from a real situation, there would probably exist other concepts, which is what happens in [Microsoft]. For example, as shown in Exhibit 11, the classification might include products such as printers, which depend on the node “Manufacturer”, which however do not fit in the category “Application”.
4. The classification implements multiple inheritance, as the authors mention “subcategories or product inherit both the definition and any assigned values from their categories”. In these circumstances, it is not possible to interpret that the nodes “Application”, “Catalog”, and “Manufacturer” are criteria, being on the other hand category-nodes.
5. Two nodes that might be criteria, “Application” and “Catalog” are parents of nodes that are not categories, these nodes apparently being products (“Programming”, “Typing”, “Desktop Publishing”, “Word processing”)
6. Given the fact that the authors are patenting an enhanced browser to be used with classifications, if they had had the intention of showing the innovative aspects that are presented in the invention of this patent application, they would have mentioned them, but they do not do it.

Exhibit 9

Application

Data management

.XXX.

5

Simulation

.XXX.

Accountancy

.XXX.

10 **Exhibit 10**

Products

For the beach

.sunglasses.

For practicing sports

15

. sunglasses.

For eye care

. sunglasses.

Exhibit 11

20 Application

Accountancy

.ABC-123.

.XYZ-890.

.Programming.

25

.Typing.

Catalog

.Desktop Publishing.

Spreadsheet

.ABC-123.

30

.XYZ-890.

.Word Processing.

Manufacturer

Company A

.ABC-123.

35

.Printer UVW.

Company B

.XYZ-890.

5 The last example of a classification that contains different types of mixed categories is
[Barnes and Noble]. In some points in the browsing process, the system shows tree fragments
that contain some similarities with the present invention, such as in Exhibit 12. However, this
proposal is far from the present invention, because as is the case with other proposals, it contains
criteria and categories which are mixed, and the categories vary as the search progresses. For
10 example, at startup there are two different categories “Business” and “History”, and later in the
process there exists a different category called “Business History”.

In order to better comprehend the difference between this system and the approach
proposed by the current patent application, Exhibit 14 shows how this search classification
would be structured if it had been created along the lines of the current invention.

15

Exhibit 12

Fiction

Fiction and literature

Graphic novels

20 Horror

Mystery and crime

Other ways to search

Audiobooks

Spanish

25 Sale

Recommended

Large Print

Exhibit 13

30 Formats

Hard cover

Soft cover

Soft cover special

Audio

35 Large Print

Exhibit 14According to origin

Fiction

5 Non Fiction

According to content

Usually fiction

Horror

Mystery and crime

10 Romance

Thriller

Usually non fiction

Business

Accounting

15 Business and commercial legislation

Business history

Africa

Gastronomy, cuisine and wine.

History

20 According to format

Paper

Audio

e-Book

25 ASSESSMENT OF THE NOVELTY AND INVENTIVE STEP OF THE INVENTION

The explanation given in the previous section shows the advantages of the invention. It has also shown several proposals that exist in the prior art that share some characteristics with the present invention, but which are nonetheless different.

Despite the fact that many of the features of the present invention are included in other
30 proposals, no proposal contains all the features simultaneously. Each one of the classification systems that have been shown presents certain problems that the present invention solves by the grouping all those features and by adding some more.

The proposals that were presented have existed for some time already. [Benson et al
1988] is from 1988. [Microsoft] existed before 2000. [Royal Academy of Spanish Language]
35 existed before 2002. The patent [IBM 2000] was filed in 1996.

The fact that so long time has passed without the appearance of a proposal like the present invention proves the inventive nature of it.

DESCRIPTION OF THE FIGURES

5 Figure 1 shows a diagram like the ones used in systemic linguistics.

Figure 2 shows a block diagram of the preferred embodiment.

Figure 3 shows a schematic example of the look of the preferred embodiment for a classification fragment.

Figure 4 shows a block scheme of an alternative embodiment.

10

EXPOSITION OF AN EMBODIMENT OF THE INVENTION

DESCRIPTION OF THE PREFERRED EMBODIMENT

15 In the preferred embodiment, the invention is built on a computerized system, which can be based, for example, on the personal computer Dell® Dimension XPS®, to which a mouse and a keyboard have been added for the user to interact with the system. In the computerized system there exists an operating system that might be, for example, Microsoft® Windows 2000®.

Figure 2 shows a block diagram of the preferred embodiment, in which the following components can be seen: a screen 2001 to observe the performance of the invention; a processing
20 unit 2002 that produces the functionality of the invention; some interaction means 2003, which would be for example a mouse, a keyboard, an optical pen or other means; and some data 2004 that contain the categories, criteria and instances that are being classified by the invention.

Additionally, the invention uses a computer tree control, such as for example the Microsoft TreeView® control. Figure 3 schematically shows how an arboreal structure could be
25 created according to the current invention for a fragment of the classification of Exhibit 1.

In the preferred embodiment, the following means are used to distinguish the criterion-nodes from the other nodes.

1. A folder icon, with a mark in the center
- 30 2. The node text starts with "According to..."
3. Red font text (which in this document is replaced by underlined font in Figure 2)

The invention is used to perform queries upon a set of instances which are categorized. In order to do that, it is first necessary to have categorized those instances, i.e. to have assigned

the categories to which the instances belong within the different criteria. In the preferred embodiment, two special methods are used in order to facilitate the categorization of instances.

The concept of DOMAIN will be presented here. A domain is a set of sister criteria that includes all the sisters of said criteria. In these circumstances, if a given instance belongs to a category that belongs to one of the criteria, it must also belong to some category of each one of the other criteria that belong to the domain. For example, in Exhibit 1, the nodes that are children of “Noun>According to nature>Event” make up a domain (the domain is composed by the nodes “According to duration” and “According to action”). If an instance has the category “Event”, it means that it must also have one or more of the categories that depend on it, which means that it must also have at least a category of each one of the criteria that belong to this domain (“According to Duration” and “According to Action”)

In these circumstances, an incomplete criterion is a criterion for which no category has been selected, even though at least one should have been selected. A complete criterion is a criterion for which the minimum number of categories have been selected. A neutral criterion is a criterion for which it is not necessary to select any category, and for which in fact there exists no selected category.

For example, in Exhibit 1, if the word “hammer” is being categorized and the category “According to nature>Entity” has been selected, the user must also select a category that belongs to the criterion “According to meaning”, because those criteria belong to the same domain. However, it is not necessary to select any category belonging to the criteria “According to duration” or “According to action”. On the other hand, if the selected category was “According to nature>Event>According to duration>Punctual”, it would be necessary to select at least a category that belongs to the criterion “According to action”, because they would belong to the same domain.

25

The method for categorizing instances comprises the following steps:

1. Selecting the instance that is to be categorized
2. Selecting a set of category nodes in the tree, with the purpose of indicating the categories to which the instance will belong
- 30 3. Identifying the complete, incomplete and neutral criteria (this step would be carried out by the invention).
4. Marking with graphical means the complete, incomplete and neutral criteria (this step is also carried out by the invention) with the purpose of facilitating the user to evaluate the current selection. In the preferred embodiment, complete criteria are marked with green background

color, incomplete criteria with red background color (and white foreground color) and neutral criteria are not marked.

The method for performing searches is carried out as explained below. The user must
5 select a set of categories, and the invention will search the instances that correspond to those categories. In this case, it is possible to leave some criteria as incomplete. If a criterion is left incomplete, such as for example "According to nature", the system will not use the categories belonging to said criterion for performing the search.

Usually, more than one category will be selected. In these circumstances, it will be
10 necessary to specify the boolean relations that must be applied, unless they are implicitly defined. For example, in Exhibit 14, if the user selects "Horror" and "Thriller", it will be necessary to specify whether he/she means "Horror AND Thriller", "Horror OR Thriller" or other boolean combination.

In the preferred embodiment, there exists a single database for each object type (for
15 example a database for words, a database for books, etc) and in each database there exist two tables. One of the tables is used to store instances, and the other one is used to store categories and criteria. In both tables, the database system assigns correlative numeric codes to the entities that are created, (instances, categories or criteria). In order to create the classifications of the instances, hyphens are used around the codes of the categories to which the instance belongs,
20 such as for example in "-1-23-22-".

DESCRIPTION OF OTHER EMBODIMENTS

It is possible to create other embodiments with a different choice of components for the computerized system, such as for example a different computer, a different tree control, a
25 different operating system, or a different element in general.

So far, it has been assumed there were three types of nodes (criterion-nodes, category-nodes and instance-nodes). In other embodiments there might be more types of nodes. For example, it is possible to also use a superhierarchy-node that might add specific properties for the characteristics that depend on it.

30 Figure 4 shows another possible embodiment of the invention, which comprises a processing unit 4001 that executes a program with the capacity to organize entities in the manner explained in this invention. This would be the case, for example for a company that is providing a data access service through Internet, to which the user would remotely access by personal computers.

In this embodiment, the invention can be used via an independent computerized system 4002, to which the invention is linked by a telecommunication system 4003. The data that are managed by the unit 4001 are integrated with the unit 4001, or they might be distributed, such as for example are the data 4005, 4006, 4007, 4008, to which the unit 4001 would link by a
5 telecommunications system 4004.

In general, the most useful arboreal structures are of the tower type, which are characterized by the fact that the different nodes are located ones on top of the others, and the nodes are differentiated, mainly, by the indentation level. The Exhibits that are shown in this document and the Microsoft Treeview® control are examples of structures of the tower type.
10 These structures are much easier to use than the ones that are used, for example, in systemic linguistics, such as the one that is shown in Figure 1.

In addition to the embodiments that are based on controls such as the Microsoft Treeview® control, it is possible to create arboreal structures by using text controls and placing them in a vertical fashion, and applying different indentation levels to the different text controls.
15 An example of these structures are the ones that are created in Internet pages by using HTML language, and it would be very similar to the structures that are shown in the Exhibits of this document.

In other embodiments it is possible to create arboreal structures that do not comprise the functionality for expanding and collapsing nodes, but they are permanently expanded. In this
20 case, the main advantage of the invention is the separation of criteria and category and the methods to manage searches and categorization.

It is possible also to implement the invention with different designs of arboreal structures. One of these designs is shown in Exhibit 15. In this arboreal structure, the level of the criterion nodes is not higher than the level of the categories that they directly dominate, but they
25 are simple differentiated by the text and format, but they have the same level of indentation. This design of arboreal structure facilitates to see the relation between categories with their parent categories, such as can be seen for example when inspecting “Noun” and “Entity”, where it is clear that “Entity” is a category that depends directly on “Noun”. In this arboreal structure, a criterion can be expanded or collapsed, and the result would be that the categories that depend on
30 it would appear or disappear without making the criterion-node itself disappear. For example, if the criterion “According to nature” is collapsed, the result would be as shown in Exhibit 16.

Exhibit 15

Words

35 Noun

According to nature

Entity

.hammer.

.brother.

5

.writer.

.cherry.

Attribute

.height.

.honesty.

10

Event

According to duration

Punctual

.arrival.

Durative

15

.concert.

.storm.

According to action

Action

.concert.

20

.arrival.

Non action

.storm.

Other

.meter.

25

.field.

According to meaning

Has utilization

.hammer.

Has function

30

.writer.

Has relationship

.brother.

Other

.height.

35

Verb

Adjective

Adverb

Closed Class

5 Exhibit 16

Words

Noun

According to nature

According to meaning

10 Has utilization

.hammer.

Has function

.writer.

Has relationship

15 .brother.

Other

.height.

Verb

Adjective

20 Adverb

Closed Class